PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H04N 7/26

A1

(11) International Publication Number: WO 97/15146

(43) International Publication Date: 24 April 1997 (24.04.97)

(21) International Application Number: PCT/IB96/01099

(22) International Filing Date: 17 October 1996 (17.10.96)

(30) Priority Data:
95202819.9 18 October 1995 (18.10.95) EP
(34) Countries for which the regional or international application was filed: NL et al.

(81) Designated States: JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(71) Applicant *(for all designated States except US)*: PHILIPS ELECTRONICS N.V. [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(71) Applicant *(for SE only)*: PHILIPS NORDEN AB [SE/SE]; Kottbygatan 7, Kista, S-164 85 Stockholm (SE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BEUKER, Rob, Anne [NL/NL]; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). THEUNIS, Hendrik, Gemmualdus, Jacobus [NL/NL]; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). HEUSDENS, Richard [NL/NL]; Prof. Holstlaan 6, NL-5656 Eindhoven (NL).

(74) Agent: SCHMITZ, Herman, J., R.; Internationaal Octrooibureau B.V., P.O. Box 220, NL-5600 AE Eindhoven (NL).

(54) Title: METHOD OF ENCODING VIDEO IMAGES

(57) Abstract

A method of encoding video images is disclosed, in which different coding methods are applied to different regions of the image. The image is divided into blocks, and the coding method which is optimal in a rate-distortion sense is selected (2) for each block. In an embodiment, transform coding (3), such as DCT or LOT, is applied to all blocks. The block size is selected in accordance with a rate-distortion criterion.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

Method of encoding video images.

FIELD OF THE INVENTION

The invention relates to a method of encoding video images, comprising the steps of dividing said images into blocks selecting one of a plurality of different coding methods for each of said blocks and encoding said blocks using the selected coding method to obtain coded data for each block. The invention also relates to an arrangement for carrying out said encoding method.

BACKGROUND OF THE INVENTION

A method of encoding video images as described in the opening paragraph is disclosed in European Patent Application EP-A 0 220 706. In this known method, transform coding is applied to each block, the block size being variable in response to brightness changes. The blocks are subdivided into smaller blocks so that the mean distortion inside each block does not exceed an allowable value.

OBJECT AND SUMMARY OF THE INVENTION

It is an object of the invention to further improve the video image encoding method.

To this end, the method according to the invention is characterized in that the step of selecting the encoding method comprises the determination of that coding method which is optimal in a rate-distortion sense. An optimal compromise between rate and distortion is thereby achieved.

In an embodiment of the method, the plurality of different coding methods is applied to pixel blocks of equal size. Examples of different coding methods are transform coding and fractal coding. In a further embodiment, the coding methods are all picture transforms, but they are applied to pixel blocks of different block sizes. Transforms used in transform coding are the Discrete Cosine Transform (DCT), the Hadamard transform, the Lapped Orthogonal Transforms (LOT), in particular the Modified LOT (MLOT), all known in the art.

In a preferred embodiment of the method it is assumed that the statistics of the image to be coded are Gaussian, and that the transform coefficients are uncorrelated. In this embodiment, the rate and distortion, on which the selection of the optimal transform type is based, can easily be calculated.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 shows a diagram of a video encoding and transmitting station employing the method according to the invention.

Fig.2 shows examples of rate-distortion curves associated with different coding methods.

Fig.3 shows a flow chart of steps carried out by a segmentation circuit which is shown in Fig.1.

Fig.4 shows a segmentation map of an image indicating the different coding methods applied to different regions of the image.

DESCRIPTION OF PREFERRED EMBODIMENTS

Fig.1 shows a diagram of a video encoding and transmitting station employing the method according to the invention. The arrangement receives a video input signal X_{in}. In an optional subtracting circuit **1**, a predicted video signal X_{pred} is subtracted therefrom. The encoder can thus operate in an intraframe mode or a (possibly motion-compensated) interframe mode. The picture to be coded is applied to a segmentation circuit **2** and a transform circuit **3**. The segmentation circuit determines, for example in a pre-analysis phase, which transform for a given block is optimal in a rate-distortion sense. The circuit further merges the contiguous blocks subjected to the same transform so as to form regions with the same transform. A "segmentation map" thus created is encoded for transmission or storage by an encoding circuit **4**.

The segmentation map is further applied to transform circuit **3** so as to indicate which transform is to be carried out during the actual coding phase. The transform coefficients obtained from transform circuit **3** are quantized and lossless coded by a quantizer and entropy coder **5**. Quantization and entropy coding are well-known in the art. For example MPEG2-like coding can be used. The coefficients for each transform block are zigzag-scanned. The DC coefficients are quantized using a fixed step size, and encoded differentially. The AC coefficients are adaptively quantized and entropy-coded using a combination of Huffman coding and run-length coding. An end-of-block code is transmitted

after the last non-zero AC coefficient of a block. The coded data thus obtained is multiplexed with the encoded segmentation map by a multiplexer **6** and transmitted to a decoder or stored on a storage medium (not shown).

The segmentation circuit **2** determines the optimal coding method in a rate-distortion sense. The rate-distortion curve of a given coding method is the collection of rate-distortion pairs (R,D) for different values of an encoding parameter t, e.g. the quantization step size of a transform coder. Fig.2 shows a rate-distortion curve **201** associated with a first coding method T1 and a second rate-distortion curve **202** associated with a second coding method T2. In the following embodiment, transform coding is applied to pixel blocks of non-equal size. The segmentation circuit **2** determines the optimal block size. In the present example, two assumptions are made to speed up the segmentation process: the statistics of the image to be coded are Gaussian, and the transform coefficients are statistically independent. Under these assumptions, the following applies (see Toby Berger: Rate Distortion Theory, A Mathematical Basis For Data Compression, Prentice-hall, Inc. Englewood Cliffs, New Jersey, 1971, pp.110-111):

1. For each pixel block k which is processed, the rate $R_k(t)$ and distortion $D_k(t)$ is:

$$R_k(t) = \frac{1}{2} \sum_i \max(\log \frac{c_{i,k}^2}{t}, 0) \quad (1)$$

$$D_k(t) = \sum_i \min(c_{i,k}^2, t) \quad (2)$$

where $c_{i,k}$ is the i-th coefficient of transform block k and t is an encoding parameter, e.g. representative of a quantizer step size.

2. The slope s of the rate-distortion curve is:

$$s = -\frac{1}{2t} \quad (3)$$

Fig.3 shows a flow chart of steps carried out by segmentation circuit **2**. In a step **21**, the circuit calculates the operating value of t in such a way that the global rate R(t) equals a required rate R_{ref}, i.e. such that:

$$R(t) = \sum_k R_k(t) = R_{ref}$$

The value of t is found, for example, by using a bi-section algorithm. Table I shows an example of such a bi-section algoritm in a pseudo-programming language. Of course, more efficient algorithms, such as Gradient methods, can be used.

Table I

t_l = minimum non-zero value of $c_{i,k}^2$;
$R_l = R(t_l)$;
t_r = maximum value of $c_{i,k}^2$;
$R_r = R(t_r)$;
repeat
    $R = R((t_l+t_r)/2)$;
    if $R > R_{ref}$
    then $t_l = (t_l+t_r)/2$;
    else $t_r = (t_l+t_r)/2$;
until $R \approx R_{ref}$
$t = (t_l+t_r)/2$

In a step 22, the circuit subjects each pixel block k to a given transform so as to obtain transform coefficients $c_{i,k}$, and calculates the rate $R_k(t)$ and distortion $D_k(t)$ for said block in accordance with equations (1) and (2), using the value t which was found in step 21. The step 22 is repeated for different block sizes. In the present example, four different transforms are considered: a 2*2 transform T1, a 4*4 transform T2, an 8*8 transform T3, or a 16*16 transform T4. In a step 23, it is checked whether or not all these transforms have been processed.

If the rate-distortion pair (R,D) has been calculated for each transform type, the best transform is selected in a step 24. The best transform is the transform for which the "Lagrangian cost" L, defined as L=R+s.D, is minimal. Herein, s is the slope of the rate-distortion curve in accordance with equation (3). An adequate way of selecting the best transform is achieved by pair-wise comparing the above transform results, i.e by carrying out the following substeps:

1. Compare, for a 4*4 block, four 2*2 T1 transform blocks with the corresponding 4*4 T2 transform block.
2. Compare, for a 8*8 block, the 8*8 T3 transform with the transform resulting from substep 1 for this block.
3. Compare, for a 16*16 block, the 16*16 T4 transform with the transform resulting from substep 2 for this block.

In a step 25, the selected transform type is stored in the segmentation map, which defines a grid determined by the smallest block size. Fig.4 shows an illustrative example of such a segmentation map.

Returning now to Fig.1, the segmentation map is applied to transform circuit 3 so as to indicate which transform type is to be used during the phase of really

encoding the image. During this encoding process, the rate $R_k(t)$ for block k as determined in step 22 may be applied to a bitrate regulation circuit (not shown in Fig.1) so as to actually achieve the rate as determined by the segmentation circuit 2. Bitrate regulation circuits are known in the art. The segmentation map is further applied to encoding circuit 4 for transmission to the decoder or storage on a storage medium. A practical encoding strategy is to assign a unique number to the different transform types. The transform number is lossless encoded, using DPCM. The resultant differences are transmitted by a combination of Huffman coding and run-length coding.

An alternative embodiment for calculating the rate-distortion pairs (step 22 above) is to actually encode (transform, quantize, Huffman and run-length coding) each potential image block k. In that case, the above assumptions (the statistics of the image to be coded are Gaussian, and the transform coefficients are uncorrelated) are not applicable.

It is also to be noted that different transforms with equal block sizes can be used in the automatic segmentation, for example Discrete Cosine Transforms, Hadamard transforms, or Lapped Transforms such as the Modified Lapped Orthogonal Transform.

It is further to be noted that a provision in the coding process is required to switch between the different transforms at the contour between regions, while maintaining (near) perfect reconstruction. For example, using linear phase transforms, this can be accomplished by mirroring at the region boundaries.

In summary, a method of encoding video images is disclosed in which different coding methods are applied to different regions of the image. The image is divided into blocks, and for each block the coding method is selected which is optimal in a rate-distortion sense. In an embodiment, transform coding, such as DCT or LOT, is applied to all blocks. The block size is selected in accordance with a rate-distortion criterion.

Claims

1. A method of encoding video images, comprising the steps of dividing said images into blocks, selecting one of a plurality of different coding methods for each of said blocks, encoding said blocks using the selected coding method to obtain coded data for each block, and transmitting data indicating the selected coding method and said coded data, characterized in that the step of selecting the encoding method comprises the determination of that coding method which is optimal in a rate-distortion sense.

2. A method as claimed in Claim 1, wherein the plurality of different coding methods is applied to pixel blocks of equal size.

3. A method as claimed in Claim 1, wherein the plurality of different coding methods are signal transforms applied to pixel blocks of different block sizes.

4. A method as claimed in Claim 3, wherein the step of determining the optimal coding method implies the calculation of the rate R(t) and distortion D(t) in accordance with

$$R_k(t) = \frac{1}{2} \sum_i \max(\log \frac{c_{i,k}^2}{t}, 0)$$

$$D_k(t) = \sum_i \min(c_{i,k}^2, t)$$

where $c_{i,k}$ is the i-th coefficient of transform block k and t is a quantization parameter.

5. An arrangement for encoding video images, comprising means for dividing said images into blocks, means for selecting one of a plurality of different coding methods for each of said blocks, means for encoding said blocks using the selected coding method for to obtain coded data for each block, and means for transmitting data indicating the selected coding method and said coded data, characterized in that the means for selecting the encoding method comprise means for determining which coding method is optimal in a rate-distortion sense.

6. An arrangement as claimed in Claim 5, wherein the plurality of different coding methods is applied to pixel blocks of equal size.

7. An arrangement as claimed in Claim 5, wherein the plurality of different coding methods are signal transforms applied to pixel blocks of different block sizes.

8. An arrangement as claimed in Claim 7, wherein means for determining the optimal coding method is adapted to calculate the rate R(t) and distortion D(t) in accordance with

$$R_k(t) = \sum_i \max(\log \frac{c_{i,k}^2}{t}, 0)$$

$$D_k(t) = \sum_i \min(c_{i,k}^2, t)$$

where $c_{i,k}$ is the i-th coefficient of transform block k and t is a quantization parameter.

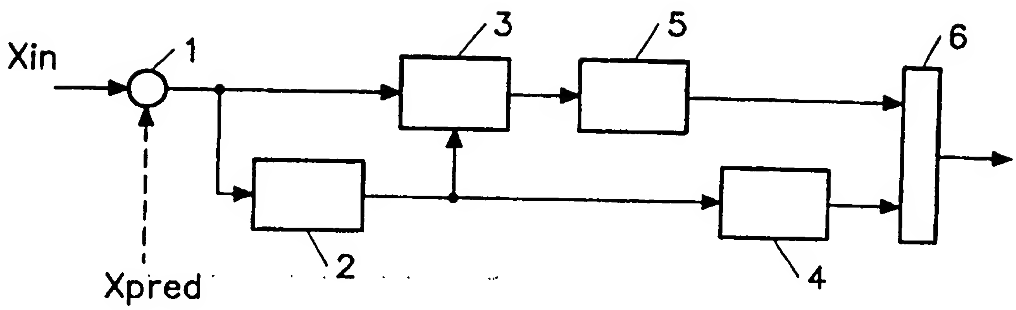

FIG. 1

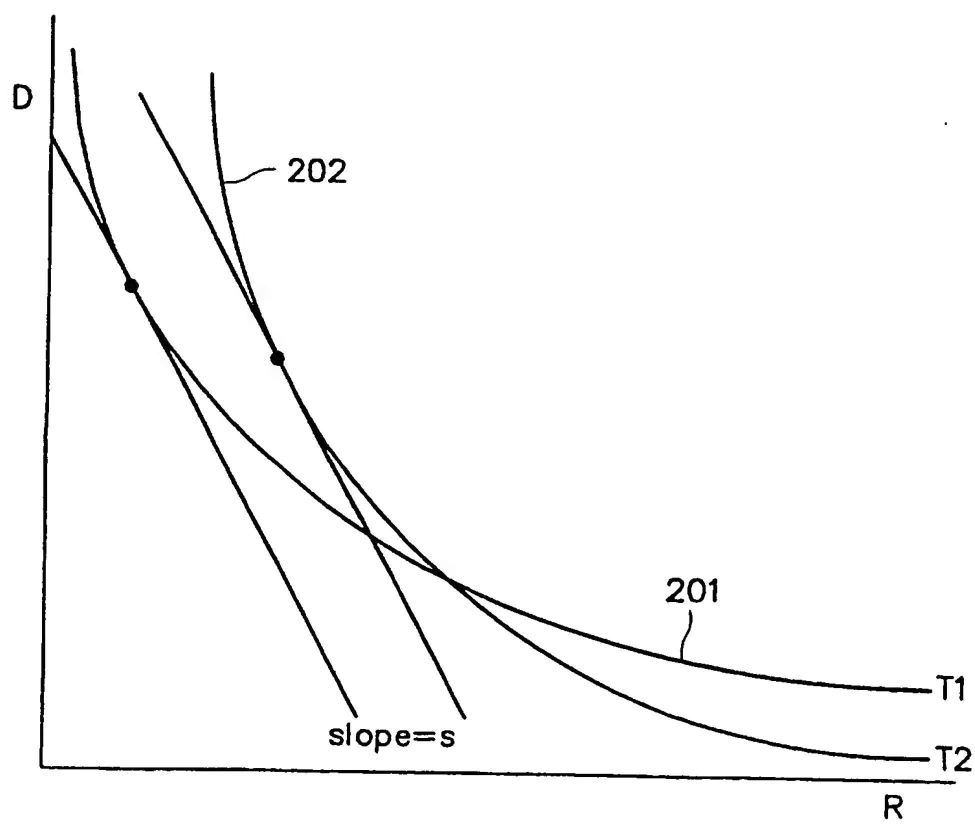

FIG. 2

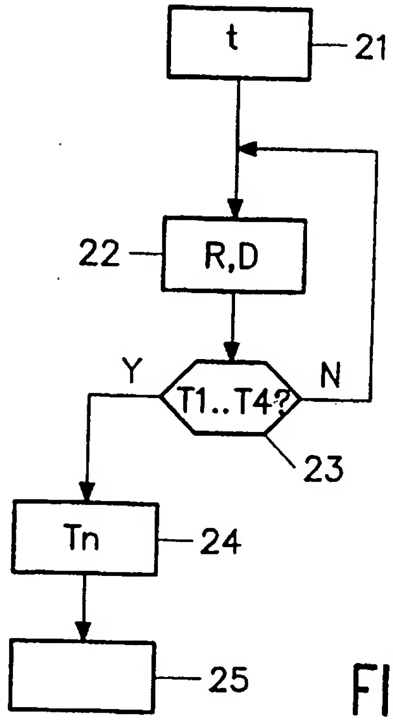

FIG. 3

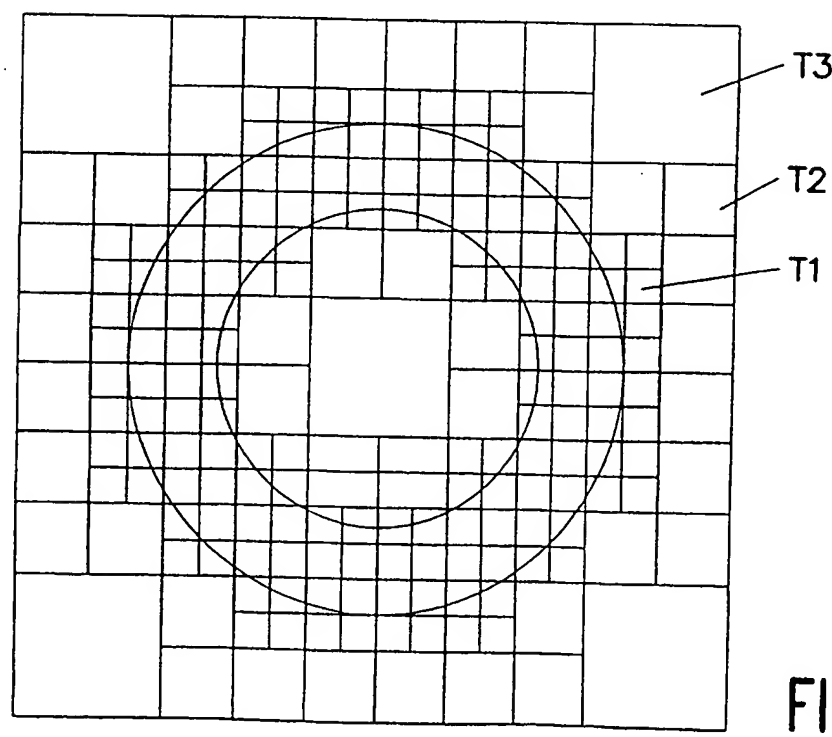

FIG. 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB 96/01099

A. CLASSIFICATION OF SUBJECT MATTER

IPC6: H04N 7/26

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC6: H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5241395 A (CHENG-TIE CHEN), 31 August 1993 (31.08.93), column 1, line 63 - column 2, line 35; column 3, line 12 - line 14 --	1-8
Y	EP 0549813 A1 (SONY CORPORATION), 7 July 1993 (07.07.93), column 1, line 10 - column 2, line 49 --	1-8
A	Toby Berger, "Rate Distortion Theory", 1971, Prentice-Hall, (New Jersey, USA), page 110 - page 111 --	4,8

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" erlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 March 1997	02 -04- 1997
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Anders Ströbeck Telephone No. +46 8 782 25 00

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB 96/01099

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	IEEE Global Telecommunications Conference Globecom"91, Volume 1, December 1991, (USA), Sullivan et al, "Rate-distortion optimized motion compensation for video compression using fixed or variable size blocks" --	3,4,7,8
A	EP 0220706 A2 (HITACHI, LTD), 6 May 1987 (06.05.87), page 3, line 26 - page 4, line 10 --	1-8
P,A	US 5506686 A (AUYEUNG ET AL), 9 April 1996 (09.04.96), column 2, line 56 - column 3, line 35 --	1-8

INTERNATIONAL SEARCH REPORT
Information on patent family members

04/03/97

International application No.
PCT/IB 96/01099

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5241395 A	31/08/93	NONE	
EP 0549813 A1	07/07/93	AU 656215 B	27/01/95
		AU 2330492 A	23/02/93
		CA 2091579 A	20/01/93
		JP 5276500 A	22/10/93
		JP 5276506 A	22/10/93
		US 5543843 A	06/08/96
		WO 9302528 A	04/02/93
EP 0220706 A2	06/05/87	DE 3686754 A	22/10/92
		JP 8024341 B	06/03/96
		JP 62101183 A	11/05/87
		US 4831659 A	16/05/89
US 5506686 A	09/04/96	AU 3506195 A	17/06/96
		CA 2178943 A	30/05/96
		EP 0740882 A	06/11/96
		WO 9616508 A	30/05/96